UNITED STATES PATENT AND TRADEMARK OFFICE

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/579,235	05/30/2006	Kazuhiro Nakadai	62533.00047	8946

32294   7590   03/25/2011
Squire, Sanders & Dempsey (US) LLP
8000 TOWERS CRESCENT DRIVE
14TH FLOOR
VIENNA, VA 22182-6212

EXAMINER
KAZEMINEZHAD, FARZAD

ART UNIT	PAPER NUMBER
2626	

NOTIFICATION DATE	DELIVERY MODE
03/25/2011	ELECTRONIC

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Notice of the Office communication was sent electronically on above-indicated "Notification Date" to the following e-mail address(es):

IPGENERALTYC@SSD.COM
SWHITNEY@SSD.COM

PTOL-90A (Rev. 04/07)

<table>
<tr><td rowspan="2">Office Action Summary</td><td>Application No.
10/579,235</td><td>Applicant(s)
NAKADAI ET AL.</td></tr>
<tr><td>Examiner
FARZAD KAZEMINEZHAD</td><td>Art Unit
2626</td></tr>
</table>

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE <u>3</u> MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1) ☒ Responsive to communication(s) filed on *12 January 2011*.

2a) ☒ This action is **FINAL**.　　　　2b) ☐ This action is non-final.

3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4) ☒ Claim(s) *1-12* is/are pending in the application.

　　4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5) ☐ Claim(s) _____ is/are allowed.

6) ☒ Claim(s) *1-12* is/are rejected.

7) ☐ Claim(s) _____ is/are objected to.

8) ☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9) ☐ The specification is objected to by the Examiner.

10) ☐ The drawing(s) filed on _____ is/are: a)☐ accepted or b)☐ objected to by the Examiner.

　　Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

　　Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11) ☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12) ☒ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

　　a)☐ All　b)☐ Some * c)☐ None of:

　　　1. ☒ Certified copies of the priority documents have been received.

　　　2. ☐ Certified copies of the priority documents have been received in Application No. _____.

　　　3. ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

　　* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)

2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)

3) ☐ Information Disclosure Statement(s) (PTO/SB/08)
　　Paper No(s)/Mail Date _____.

4) ☐ Interview Summary (PTO-413)
　　Paper No(s)/Mail Date _____.

5) ☐ Notice of Informal Patent Application

6) ☐ Other: _____.

DETAILED ACTION

Response to Amendment

1.   In response to the office action from 10/15/2010, the applicant has filed an

amendment, filed 1/12/2011, traversing the previous rejection "as being based on an

improper prior art reference", Kim (US 2004/0175006).

2.   According to **MPEP 706.02(b)(E)**: "The filing date of the priority document is not

perfected unless applicant has filed a certified priority document in the application **(and**

an English language translation, if the document is not in English) (see 37 CFR

1.55(a)(3)) and the examiner has established that the priority document satisfies the

enablement and description requirements of 35 U.S.C. 112, first paragraph". The

English language translation of the foreign priority was received on 1/12/2011, while the

previous office action was dated October 15, 2010. Therefore at the time of the previous

examination, the foreign priority was not yet perfected, and the Kim reference was a

proper prior art reference.

3.   The applicant is therefore respectfully directed to the new office action further in

view of Ichikawa et al. (US Patent 7,478,041).

Claim Rejections - 35 USC § 103

4.   The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all

obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set
> forth in section 102 of this title, if the differences between the subject matter sought to be patented and
> the prior art are such that the subject matter as a whole would have been obvious at the time the
> invention was made to a person having ordinary skill in the art to which said subject matter pertains.
> Patentability shall not be negatived by the manner in which the invention was made.

5.   Claims 1, 4, 6-7 are rejected under 35 U.S.C. 103(a) as being unpatentable over

Asano (US 2004/0054531), and further in view of Ichikawa et al. (US Patent 7,478,041).

Regarding claim 1, Asano does teach an automatic speech recognition system, which recognizes speeches in acoustic signals detected by a plurality of microphones as character information, the system comprising:

a sound source localization module configured to localize a sound direction corresponding to a specified speaker based on the acoustic signals detected by the plurality of microphones (¶ 0129 teach the head unit 3 in Fig. 2 (performing similar function as the sound source localization module) enables obtaining the direction of sound utilizing a plurality of microphones at a target (e.g. a robot) which receives the speech signals by computing power and phase differences of speech signals due to sources attributed to users (speakers); Abstract teaches all incoming speech undergo speech recognition utilizing plural sets of acoustic models);

a feature extractor configured to extract features of speech signals contained in one or more pieces of information detected by the plurality of microphones (the feature extractor unit 101 in Fig. 9 receives speech data from microphones (unit 21 Fig. 9) via the analog to digital converter. ¶ 0104 lines 1-4 teach extracting feature vectors of the incoming speech data) ;

an acoustic model memory configured to store distance-dependent acoustic models that are adjusted to a plurality of distances at intervals (¶ 0131 lines 4-8 referring to Fig. 9 disclose storing acoustic models corresponding to selected distances in the database units (104)_1 …. (104)_N (located in the memory module 42 in Fig. 3);

these data are used following determination of distance of a sound source attributed to

a user; ¶ 0106 teaches all feature vectors corresponding to acoustic analysis (¶0104

lines 1-2) are stored in speech periods (intervals) corresponding to an utterance);

an acoustic model composition module configured to compose an acoustic model

adjusted to the sound distance, which is localized by the sound source localization

module, based on the distance-dependent acoustic models in the acoustic model

memory, the acoustic model composition module also configured to store the acoustic

model in the acoustic model memory (¶ 0114 teaches "N" acoustic models

corresponding to "N" sound sources located at "N" distances are produced (composed)

where each acoustic model corresponding to a certain distance (localization) is stored in

a certain database (e.g. one of the units (104)_1 to (104)_N) in Fig. 9));

and a speech recognition module configured to recognize the features extracted

by the feature extractor as character information using the acoustic model composed by

the acoustic model composition module (¶ 0132 (lines 2-1 above ¶ 0133) teaches

module units 41B and 41A in Fig. 3 as disclosed in step S4 in Fig. 10 enable

performing speech recognition utilizing feature vectors extracted from the speech data

(¶ 0132 lines 1-3); ¶ 0108 lines 1-4 teach acoustic model databases used for the

speech recognition have stored acoustic characteristics of phonetic-linguistic-units such

as phonemes or syllables (e.g. character units comprising words) in them).

Asano does not specifically disclose:

Wherein the acoustic model composition module is configured to compose an

acoustic model adjusted to the sound direction.

Ichikawa et al. does teach an acoustic model composition module configured to

compose an acoustic model adjusted to the sound direction (Using the profile fitting unit

33 in Fig. 3 (Acoustic model composition module) mentioned in Col. 10 lines 59-60,

according to Col. 10 lines 61-67 "by using a base form sound for various frequencies"

(acoustic unit corresponding to acoustic feature power P as a function of the frequency),

"profile P(theta) (acoustic unit adjusted to the sound direction (see next)) of the

microphone array 111(Fig. 1) .... is obtained (adjusts the power feature or power

acoustic unit) beforehand in possible various sound directions" ; here the profiles

P(theta) are stored in database 50 (Fig. 2) according to Col. 10 lines 6-9, which are

used by the sound source localization part unit 20 (sound source localization module)

according to Col. 9 lines 27-29 and "for estimating a sound source direction of the

voice" (localizing the sound source of the input voice) according to Col. 3 lines 46-47; P

and P(theta) are identified as the "voice power distribution data" (Col. 10 line 43-44)

where power is a well know acoustic feature; "theta" is the angle signifying the direction

of the sound).

It would have therefore been obvious to one with ordinary skill in the art at the

time the invention was made to incorporate modules 33 in Fig. 3 of Ichikawa et al. into

the feature extractor module 101 of Fig. 9 of Asano would enable the combined

modules to function in combination as they do separately and to further enable Asano to

not only generate and store distance-dependent acoustic models in the memory as is

done here in modules 104 in Fig. 9, but also direction-dependent acoustic models

obtained by using the newly determined direction dependent acoustic features, such as

power, so that the robot will have superior bias for detecting sound direction by adjusting the power or volume of incoming sound frequencies as a function of angular direction and will point at the direction of the source associated with the sound frequency to reduce detecting wrong signals and enhance its speech recognition performance.

Regarding claim 4, Asano does teach a system according to claim 1, wherein the sound source localization module is further configured to employ a scattering theory to generate a model for an acoustic signal, which scatters on a surface of a member (¶ 0142 lines 4-9 and ¶ 0145 teach using reflected (scattered) ultrasonic wave pulses from an obstacle, the robot can determine the distance of a user producing speech from the robot which according to ¶ 0131 is used in generating distance dependent acoustic models; all these processes are made possible by the sensor unit 111 (Fig. 11) attached to the robot);

specifying the sound direction for the speaker with the intensity difference and the phase difference detected from the plurality of microphones (¶ 0129 lines 1-5);

Asano does not specifically teach using scattered (reflected) waves from the surface to which the microphones are attached to determine the sound direction attributed to a user (speaker). But it would have been obvious to one with ordinary skill in the art at the time the invention was made to utilize the sensor unit 111 to create a second model in which one determines the phase and power difference of the sound waves reflected (scattered) from the surface of the robot in determining the direction of

the speech signals and thereby create a second model of determination of direction of a

sound wave incident on the robot by analyzing reflected rather than incident waves on

the robot and thereby help in generating more accurate sound direction by

benchmarking the results of the two models against one another.

Regarding claim 6, Asano does not specifically disclose a system according to

claim 1, wherein the acoustic model composition module is configured to compose an

acoustic model for the sound direction by applying weighted linear summation to the

direction- dependent acoustic models in the acoustic model memory, and weights

introduced into the linear summation are determined by training.

Ichikawa et al. does disclose a system according to claim 1, wherein the

acoustic model composition module is configured to compose an acoustic model for the

sound direction by applying weighted linear summation to the direction- dependent

acoustic models in the acoustic model memory, and weights introduced into the linear

summation are determined by training (equation 1 (Col. 11 line 15) clearly shows the

acoustic unit X(theta) obtained as a linear weighted summation of the direction

dependent voice power distribution P(theta) (power acoustic feature) and Q

(nondirectional background sound (background noise power which is another acoustic

feature)) where the said coefficients "are decided (trained) so as to minimize an

evaluation function"; i.e., Col. 11 lines 7-15 referring to equation 1 in line 15 teach

"Profile X(theta) (acoustic unit for a sound direction) obtained (composed at the module

33 (acoustic model composition module) as it follows calculation and storage of P(theta)

discussed in claim 1) for the observed voice can be approximated by a sum

(summation) of respective coefficient multiples of directional (of linear weighted) sound

source profile P(theta) for sound source from a given direction (direction dependent

power distribution (acoustic unit)), and profile Q for nondirectional background sound")

For obviousness to combine see claim 1.

Regarding claim 7, Asano does teach a system according to claim 1, further

comprising a speaker identification module, wherein the acoustic model memory is

further configured to possess the direction-dependent acoustic models for respective

speakers (identical to claim 1, 3rd limitation and rejected under similar rationale),

and wherein the acoustic model composition module is further configured to:

refer to distance-dependent acoustic models of a speaker who is identified by the

speaker identifying module and to a sound distance localized by the sound source

localization module (¶ 0116 teaches the acoustic model database to have the acoustic

model of speakers (specific speakers) at specified locations; ¶ 0159 teaches those

acoustic models are produced by from the speech data acquired by microphone placed

close to the mouth of the speakers for better voice recognition leading to better speaker

identification; ¶0130 lines 5-1 above ¶0131 teach image of the face of a user (e.g. a

potential speaker) taken by the robot's CCD cameras (modules 22L and 22R in Fig. 2)

are used as a reference pattern (stored) for image recognition (user identification);

finally ¶ 0145 lines 1-5 teach the robot uses both detection of a user uttering (identifying

a user by his voice) as well as image recognition (identifying a user by his image) in

orienting his head unit in the direction of the user; i.e., speaker identification is done by both image pattern recognition as well as by voice recognition using the stored acoustic models of the user (speaker));

compose an acoustic model for the sound direction based on the direction-dependent acoustic models in the acoustic model memory; and storing the acoustic model in the acoustic model memory (corresponds to the third limitation of the first claim and rejected under similar rationale).

For obviousness analysis please see claim 1.

1.   Claims 2-3, 8-12 are rejected under 35 U.S.C. 103(a) as being unpatentable over Asano in view of Ichikawa et al., and further in view of Ito et al. (US Patent 7,076,433).

Regarding claim 2, Asano in view of Ichikawa et al. do teach an automatic speech recognition system, which recognizes speeches of a specified speaker in acoustic signals detected by a plurality of microphones as character information, the system comprising:

a sound source localization module configured to localize a sound direction corresponding to the specified speaker based on the acoustic signals detected by the plurality of microphones (identical to the first limitation of claim 1 and rejected under similar rationale);

an acoustic model memory configured to store distance-dependent acoustic models that are adjusted to a plurality of directions at intervals (identical to the 3rd

limitation of claim 1 and rejected under similar rationale) ;

an acoustic model composition module configured to compose an acoustic model adjusted to the sound direction, which is localized by the sound source localization module, based on the distance-dependent acoustic models in the acoustic model memory, the acoustic model composition module storing the acoustic model in the acoustic model memory (identical to the 4th limitation of claim 1 and rejected under similar rationale) ;

and a speech recognition module configured to recognize the features extracted by the feature extractor as character information using the acoustic model composed by the acoustic model composition module (identical to the 5th limitation of claim 1 and rejected under similar rationale)

Asano in view of Ichikawa et al. do not specifically disclose a sound source separation module which separates speech signals of the specified speaker from the acoustic signals based on the sound direction localized by the sound source localization module

a feature extractor configured to extract features of the speech signals separated by the sound source separation module;

an acoustic model memory configured to store direction-dependent acoustic models that are adjusted to a plurality of directions at intervals corresponding to speech signals.

Ito et al. does teach a sound source separation apparatus which in one embodiment separates sound (e.g. attributed to a speaker) from a mixed input signal

(acoustic signal) by utilizing sound source direction as an acoustic feature (Abstract lines 1-2 first ¶ and line 4 from the bottom; Col. 18 lines 61 to 66 referring to Fig. 16 unit 911). Furthermore it teaches its acoustic feature extractor to incorporate a sound source direction prediction layer which aids in determination of the peaks corresponding to acoustic features of the sound source incident from a certain direction (Col. 19 lines 33-40 and module 921 in Fig. 16).

It would have therefore been obvious to one with ordinary skill in the art at the time the invention was made that utilizing the modules 915 and 921 in Fig. 16 of Ito et al. into the feature extractor unit 101 of Fig. 9 of Asano would enable Asano in view of Ichikawa et al. to separate a sound (e.g a speech signal) from a mixed input when the sound is incident from a certain direction by utilizing its direction dependent features enabling the robot to obtain bias for direction for a certain speaker and will point at the direction of the source of sound to reduce detecting wrong signals and enhance its speech recognition performance. Storage of direction dependent features also aids in avoiding calculations of the said features and helps in raising efficiency.

Regarding claim 3, Asano in view of Ichikawa et al. do suggest a system according to claim 1, wherein the sound source localization module is further configured to:

acquire an intensity difference and a phase difference for the harmonic relationships extracted through the plurality of microphones (Asano ¶ 0129 lines 1-5 teach acquiring power (intensity) and phase difference of speech signals (which

maintain harmonic spectra) picked up by a device (e.g. robot) microphones);

acquire belief factors for a sound direction based on the intensity difference and

the phase difference, respectively; and

determine a most probable sound direction (utilizing power and phase difference

in determining the direction of sound as disclosed in ¶ 0129 lines 1-5 does inherently

involve these or equivalent steps to achieve the same net outcome (sound direction

determination)).

Asano in view of Ichikawa et al. does not specifically disclose:

To perform a frequency analysis for the acoustic signals detected by the

microphones to extract harmonic relationships.

Ito et al. does disclose performing a frequency analysis for the acoustic signals

detected by the microphones to extract harmonic relationships (Col. 12 lines 45-50

teach harmonic calculation layer (named as the intermediate feature extraction layer

unit 107 in Fig. 9) determines harmonic features of features (attributed to acoustic

analysis of speech signals) at each time based on their frequency variation rates (i.e. by

doing frequency analysis));

It would have therefore been obvious to one with ordinary skill in the art at the

time the invention was made that utilizing the harmonic calculation layer (unit 107 in Fig.

9) into the feature extractor unit 101 of Fig. 9 of Asano would enable Asano in view of

Ichikawa et al. to extract harmonic structures attributed to speech signals incident on

the microphones of the robot prior to obtaining their phase and power differences for the

sake of determining the direction of the speech and thereby eliminate redundant parts of

the spectra in determining direction of speech which primarily included harmonics and

thereby enhance efficiency and accuracy.

Regarding claim 8, the preamble, the 1st, 4th, 5th, 6th and 7th limitations

correspond to the preamble, 1st, 2nd, 3rd, 4th and 5th limitations respectively of the claim

1 and are therefore rejected under similar rationale over Asano in view of Ichikawa et

al..

Limitation 3, corresponds to the limitation 2 of the claim 2 and is therefore

rejected under similar rationale by Ito et al.

Regarding the 2nd limitation, storing direction dependent acoustic models at time

intervals (the 3rd limitation of claim 1) will amount to storing the sound direction at time

intervals corresponding to the speech attributed to a sound source and the storage

medium used (i.e. the memory unit 42 in Fig. 3) will enable the unit the functionality of

the stream tracking module claimed in this limitation which will enable estimating the

current position of a sound source by simple interpolation.

For obviousness please see claims 1 and 2.

Regarding claims 9, 10, 11, 12, they correspond to claims 3, 4, 6, 7 respectively

with identical limitations and are therefore rejected under similar rationales.

6.   Claim 5 rejected under 35 U.S.C. 103(a) as being unpatentable over Asano in

view of Ichikawa et al. and Ito et al., and further in view of Okuno et al. (US Patent

7,035,418).

Regarding claim 5, Asano in view of Ichikawa et al. and Ito et al. do not specifically disclose a system according to claim 2, wherein the sound source separation module is further configured to employ an active direction-pass filter so as to separate speeches, the filter is configured to:

separate speeches by a narrower directional band when a sound direction, which is localized by the sound source localization module, lies close to a front, which is defined by an arrangement of the plurality of microphones;

and separate speeches by a wider directional band when the sound direction lies apart from the front.

Okuno et al. does teach a directional filter for sound based on the position of the source, enabling localization and extracting sound (speech) information of the said sound source which will thereby enable it to separate that source from other sound (speech) sources (Col. 4 lines 25-34); Col. 8 lines 1-7 referring to the flow chart on Fig. 8 identifies steps ST5 and ST6 as the steps associated with the directional filter's operation; Col. 8 lines 15-20 note the results of directional filter on detecting directions of sound from three sources (A, B and C on Fig. 7) and notes that the angular range (directional band) about which these directions are determined to have been reduced due to the application of the filter; Col. 11 lines 48-52 teach the directional filter functions according to the direction and position of the sound source which is determined first by computing the difference between phase and intensity of sound received at two receivers from the same source which is directly related to their path

differences; Col. 6 lines 32-35 referring to Fig. 4 explicitly show a relationship (d=D

sin(theta)) between the two "sound" signal path differences (attributed to the source

position which is called "d"), the distance between the receiving microphones (called

"D") and the angle governing the direction of the sound source with respect to an axis at

right angles to the line connecting the two microphones (called "theta") which shows the

angle and thereby the angular range (directional band) to increase with that path

difference.

It would have therefore been obvious to one with ordinary skill in the art at the

time the invention was made that utilizing these methods (i.e., steps ST5 and ST6 in

Fig. 8 of Okuno et al.) for directional filter into the flow chart of Fig. 10 of Asano (to

after step S2) would enable Asano in view of Ichikawa et al. and Ito et al. to apply the

distance dependent directional filter of Okuno et al. which would result in less accuracy

and larger angular range (wider directional band) for further sound sources since the

incoming signals possess longer path differences as larger path difference (i.e. larger

"d" in the relationship above) lead to larger angular range (i.e. larger deviations

(directional band) in "theta"). Application of directional filter will reduce the error in

determining the direction of the sound source in general.

Conclusion

The prior art made of record and not relied upon is considered pertinent to

applicant's disclosure. Botterweck (US 2002/0120444).

2.   **THIS ACTION IS MADE FINAL.** Applicant is reminded of the extension of time

policy as set forth in 37 CFR 1.136(a).

A shortened statutory period for reply to this final action is set to expire THREE

MONTHS from the mailing date of this action. In the event a first reply is filed within

TWO MONTHS of the mailing date of this final action and the advisory action is not

mailed until after the end of the THREE-MONTH shortened statutory period, then the

shortened statutory period will expire on the date the advisory action is mailed, and any

extension fee pursuant to 37 CFR 1.136(a) will be calculated from the mailing date of

the advisory action. In no event, however, will the statutory period for reply expire later

than SIX MONTHS from the mailing date of this final action.

Any inquiry concerning this communication or earlier communications from the

examiner should be directed to FARZAD KAZEMINEZHAD whose telephone number is

(571)270-5860. The examiner can normally be reached on M-F 8:30AM-5:00 PM EST.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's

supervisor, Talivaldis I. Smits can be reached on (571)272-7628. The fax phone

number for the organization where this application or proceeding is assigned is 571-

273-8300.

Information regarding the status of an application may be obtained from the

Patent Application Information Retrieval (PAIR) system. Status information for

published applications may be obtained from either Private PAIR or Public PAIR.

Status information for unpublished applications is available through Private PAIR only.

For more information about the PAIR system, see http://pair-direct.uspto.gov. Should

you have questions on access to the Private PAIR system, contact the Electronic

Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a

USPTO Customer Service Representative or access to the automated information

system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

/FK/

/Talivaldis Ivars Smits/
Primary Examiner, Art Unit 2626   03/21/2011